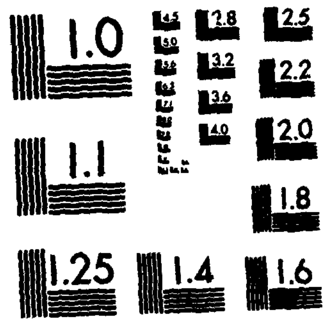
AD-A122 942 MONOTONICITY IN SELECTION PROBLEMS: A UNIFIED APPROACH 1/1
(U) FLORIDA STATE UNIV TALLAHASSEE DEPT OF STATISTICS
R L BERGER ET AL. SEP 82 FSU-STATISTICS-M647
UNCLASSIFIED ARO-19367.2-MA DAAG29-82-K-0168 F/G 12/1 NL

END
DATE
FILMED
2 83
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

ARO 19367.2-MA

(12)

MONOTONICITY IN SELECTION PROBLEMS: A UNIFIED APPROACH

by

Roger L. Berger[1] and Frank Proschan[2]

FSU Statistics Report M647
USARO Technical Report D-57
AFOSR Technical Report 82-154

AD A122942

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

September, 1982

American Mathematical Society 1980 subject classifications: Primary 62F07; Secondary 62H10.

Key words and phrases. Ranking and selection, decreasing in transposition, arrangement increasing, monotonicity, comparison with a standard.

[1]This research was supported in part by the U.S. Army Research Office Grant No. DAAG 29-82-K-0168.

[2]Research sponsored by the Air Force Office of Scientific Research Grant AFOSR 82-K-0007.

MONOTONICITY IN SELECTION PROBLEMS: A UNIFIED APPROACH

by

Roger L. Berger and Frank Proschan

ABSTRACT

Let $\underset{\sim}{X} = (X_1, \ldots, X_n)$ have a density $g(\underset{\sim}{x}, \underset{\sim}{\lambda})$ which is decreasing in transposition, where $\underset{\sim}{\lambda} = (\lambda_1, \ldots, \lambda_n)$. Suppose one wishes to select a subset of $\{1, \ldots, n\}$ containing the subscripts associated with the largest values of the λ_i's. Let $S(\underset{\sim}{x})$ be a permutation invariant selection rule which always selects a subset associated with the largest values of the X_i's. Let $A = \{i(1),\ldots,i(k)\} \subset \{1,\ldots,n\}$ and $B = \{j(1),\ldots,j(k)\} \subset \{1,\ldots,n\}$ be such that $\lambda_{i(s)} \geq \lambda_{j(s)}$, $s = 1,\ldots,k$. Then the following three inequalities are proved. ($|C|$ denotes the number of elements in C. C^c denotes the complement of C.)
(i) $P_{\underset{\sim}{\lambda}}(|A \cap S(\underset{\sim}{X})| \geq m) \geq P_{\underset{\sim}{\lambda}}(|B \cap S(\underset{\sim}{X})| \geq m)$ for every $m \in R$,
(ii) $P_{\underset{\sim}{\lambda}}(A = S(\underset{\sim}{X})) \geq P_{\underset{\sim}{\lambda}}(B = S(\underset{\sim}{X}))$, and (iii) $P_{\underset{\sim}{\lambda}}(|A^c \cap S(\underset{\sim}{X})| \leq m) \geq P_{\underset{\sim}{\lambda}}(|B^c \cap S(\underset{\sim}{X})| \leq m)$ for every $m \in R$. These generalized monotonicity properties are derived using a unified approach. The results apply to selection rules proposed under several formulations of the selection problem.

1. Introduction. In this paper we study some monotonicity properties of ranking and selection rules. Let $\underset{\sim}{X} = (X_1, \ldots, X_n)$ be a random observation with distribution $F(\underset{\sim}{x}; \underset{\sim}{\lambda})$, where the unknown parameter vector $\underset{\sim}{\lambda} = (\lambda_1, \ldots, \lambda_n) \in \Lambda \subset R^n$. The general goal of a selection problem is to decide which coordinates of $\underset{\sim}{\lambda}$ are the largest or which are larger than a value λ_0 (possibly unknown). This is accomplished by selecting $S(\underset{\sim}{x}) \subset \{1, \ldots, n\}$, depending on $\underset{\sim}{X} = \underset{\sim}{x}$, and asserting that the largest parameters are in $\{\lambda_i: \ i \in S(\underset{\sim}{x})\}$. The subset $S(\underset{\sim}{X})$ may be of fixed or random size depending on the formulation of the selection problem under consideration. See, for example, Bechhofer (1954), Gupta and Sobel (1958), Lehmann (1961), Gupta (1965) and Tong (1969) for five formulations. In this paper, we will not be concerned with a specific formulation of the selection problem or with rules which satisfy a specific probability requirement (sometimes called a P^*-condition). For our purposes a (nonrandomized) selection rule $S(\underset{\sim}{x})$ is any mapping from the sample space $\mathcal{X}$ of $\underset{\sim}{X}$ into the set of subsets of $\{1, \ldots, n\}$. The selection rules described in the above five formulations all satisfy this definition and the results derived herein apply to these rules. Gupta (1965) calls a selection rule monotone if $\lambda_i \geq \lambda_j$ implies $P_{\underset{\sim}{\lambda}}(i \in S(\underset{\sim}{X})) \geq P_{\underset{\sim}{\lambda}}(j \in S(\underset{\sim}{X}))$. This monotonicity property is a desirable property for a selection rule, given the goal of selecting a subset consisting of the large values of λ_i. Many authors have shown that their heuristically proposed selection rules are monotone. For example, Santner (1975) proved the monotonicity of a large class of selection

rules. In this paper we generalize the above notion of monotonicity and present some other notions of monotonicity. Then we show in a unified way that for many selection problems a large class of selection rules possess these monotonicity properties.

The monotonicity properties we consider are the following. Let $A = \{a_1, \ldots, a_k\}$ and $B = \{b_1, \ldots, b_k\}$ denote two subsets of $\{1, \ldots, n\}$ with $|A| = |B| = k$, where $|\cdot|$ denotes subset size. Subset A <u>is better than</u> B if, for some arrangements $a_{i(1)},\ldots,a_{i(k)}$ and $b_{j(1)},\ldots,b_{j(k)}$ of the elements of A and B, $\lambda_{a_{i(r)}} \geq \lambda_{b_{j(r)}}$ for every $r = 1, \ldots, k$. If A is better than B, then each of the following inequalities would be desirable for a selection rule:

$$P_{\underset{\sim}{\lambda}}(|A \cap S(\underset{\sim}{X})| \geq m) \geq P_{\underset{\sim}{\lambda}}(|B \cap S(\underset{\sim}{X})| \geq m) \text{ for every } m, -\infty \leq m < \infty; \quad (1.1)$$

[In words, $P_{\underset{\sim}{\lambda}}$(at least m of the elements of A are selected) $\geq P_{\underset{\sim}{\lambda}}$(at least m of the elements of B are selected).]

$$P_{\underset{\sim}{\lambda}}(A = S(\underset{\sim}{X})) \geq P_{\underset{\sim}{\lambda}}(B = S(\underset{\sim}{X})); \quad (1.2)$$

$$P_{\underset{\sim}{\lambda}}(|A^c \cap S(\underset{\sim}{X})| \leq m) \geq P_{\underset{\sim}{\lambda}}(|B^c \cap S(\underset{\sim}{X})| \leq m) \text{ for every } m, -\infty \leq m < \infty. \quad (1.3)$$

Some special cases may be of particular interest. By setting $m = k$ in (1.1) we obtain $P_{\underset{\sim}{\lambda}}(A \subset S(\underset{\sim}{X})) \geq P_{\underset{\sim}{\lambda}}(B \subset S(\underset{\sim}{X}))$. Furthermore, if $k = 1$, we obtain the classical monotonicity property of Gupta (1965). By setting $m = 0$ in (1.3) we obtain $P_{\underset{\sim}{\lambda}}(A \supset S(\underset{\sim}{X})) \geq P_{\underset{\sim}{\lambda}}(B \supset S(\underset{\sim}{X}))$.

In Section 2 we present the notation we will use and the assumptions concerning $F(\underset{\sim}{x}; \underset{\sim}{\lambda})$ and $S(\underset{\sim}{x})$ we will make. The main results, inequalities

(1.1), (1.2), and (1.3), are proved in Section 3. The extension of these results to include additional parameters and rules based on statistics other than $\underset{\sim}{X}$ is outlined in Section 4.

In a similar fashion, we can obtain dual results for the monotonicity properties of ranking and selection rules designed to select the smallest rather than the largest values of the parameter. We omit the detailed statement of these dual results and of their proofs.

2. Notation and assumptions. Let $\underset{\sim}{\pi} = (\pi_1, \ldots, \pi_n)$ denote a permutation of $(1, \ldots, n)$. For any $\underset{\sim}{x} \in R^n$, let $\underset{\sim}{x} \circ \underset{\sim}{\pi}$ denote $(x_{\pi_1}, \ldots, x_{\pi_n})$. Let $g(\underset{\sim}{x}; \underset{\sim}{\lambda})$ be a function from R^{2n} into R. We say that g is permutation invariant if $g(\underset{\sim}{x} \circ \underset{\sim}{\pi}; \underset{\sim}{\lambda} \circ \underset{\sim}{\pi}) = g(\underset{\sim}{x}; \underset{\sim}{\lambda})$ for every $\underset{\sim}{x} \in R^n$, every $\underset{\sim}{\lambda} \in R^n$, and every permutation $\underset{\sim}{\pi}$.

Let $\underset{\sim}{\pi}$ and $\underset{\sim}{\pi}'$ be two permutations such that

$$\begin{aligned} \pi &= (\pi_1, \ldots, \pi_i, \ldots, \pi_j, \ldots, \pi_n) \text{ and} \\ \pi' &= (\pi_1, \ldots, \pi_j, \ldots, \pi_i, \ldots, \pi_n), \end{aligned} \tag{2.1}$$

where $i < j$ and $\pi_i < \pi_j$. We say that $\underset{\sim}{\pi}'$ is a simple transposition of $\underset{\sim}{\pi}$; in symbols, $\underset{\sim}{\pi} >^t \underset{\sim}{\pi}'$. Let $\underset{\sim}{\pi}$ and $\underset{\sim}{\pi}'$ be such that there exists a finite number of permutations, $\underset{\sim}{\pi}^0, \underset{\sim}{\pi}^1, \ldots, \underset{\sim}{\pi}^k$, satisfying $\underset{\sim}{\pi} = \underset{\sim}{\pi}^0 >^t \underset{\sim}{\pi}^1 >^t \ldots >^t \underset{\sim}{\pi}^k = \underset{\sim}{\pi}'$. We say $\underset{\sim}{\pi}'$ is a transposition of $\underset{\sim}{\pi}$. By extension of notation, we will say $\underset{\sim}{\lambda} \circ \underset{\sim}{\pi}'$ is a transposition of $\underset{\sim}{\lambda} \circ \underset{\sim}{\pi}$ if $\underset{\sim}{\pi}'$ is a transposition of $\underset{\sim}{\pi}$ and $\lambda_1 \le \ldots \le \lambda_n$.

The concept of a decreasing in transposition function will play a central role in our exposition. The function $g(\underset{\sim}{x}; \underset{\sim}{\lambda})$ is decreasing in transposition (DT) if

$$g(\underset{\sim}{x}; \underset{\sim}{\lambda}) \text{ is permutation invariant,} \tag{2.2}$$

and

$$x_1 \le \dots \le x_n,\ \lambda_1 \le \dots \le \lambda_n, \text{ and } \underset{\sim}{\pi} \overset{t}{>} \underset{\sim}{\pi}' \tag{2.3}$$

imply

$$g(\underset{\sim}{x}, \underset{\sim}{\lambda} \circ \underset{\sim}{\pi}) \ge g(\underset{\sim}{x}; \underset{\sim}{\lambda} \circ \underset{\sim}{\pi}').$$

Hollander, Proschan, and Sethuraman (1977) (HPS(1977)) present a detailed investigation of DT functions and many examples. The DT property is called arrangement increasing by Marshall and Olkin (1979).

We assume that the observation vector $\underset{\sim}{X} = (X_1, \dots, X_n)$ has a density $g(\underset{\sim}{x}; \underset{\sim}{\lambda})$ with respect to a measure $\sigma(\underset{\sim}{x})$, where σ satisfies $\int_A d\sigma(\underset{\sim}{x}) = \int_A d\sigma(\underset{\sim}{x} \circ \underset{\sim}{\pi})$ for each permutation $\underset{\sim}{\pi}$ and Borel set $A \subset R^n$. We assume that g is DT. HPS(1977) list several discrete and continuous densities which are DT. For example, if $g(\underset{\sim}{x}; \underset{\sim}{\lambda}) = \pi_{i=1}^{n} h(x_i; \lambda_i)$ and h is TP_2 then g is DT. Eaton (1967), Hsu (1977), and Gupta and Miescke (1982) have investigated selection problems involving a DT density. Our results differ from theirs in that they compared the operating characteristics or risk functions of different selection rules, whereas, inequalities (1.1), (1.2), and (1.3) compare different operating characteristics of a single selection rule at a time.

A nonrandomized selection rule $S(\underset{\sim}{x})$ can be defined by specifying its <u>individual selection probabilities</u>, $\psi_1(\underset{\sim}{x}), \dots, \psi_n(\underset{\sim}{x})$, which are defined by

$$\psi_i(\underset{\sim}{x}) = \begin{cases} 1 & \text{if } i \in S(\underset{\sim}{x}) \\ 0 & \text{if } i \notin S(\underset{\sim}{x}) \end{cases} \tag{2.4}$$

We will make the following assumptions about $S(\underset{\sim}{x})$:

$$\text{if } \psi_i(\underset{\sim}{x}) = 1 \text{ and } x_j \geq x_i \text{ then } \psi_j(\underset{\sim}{x}) = 1; \tag{2.5}$$

and

$$\psi_{\pi_i}(\underset{\sim}{x}) = \psi_i(\underset{\sim}{x} \circ \underset{\sim}{\pi}) \text{ for every } \underset{\sim}{x} \in R^n, \text{ every } i \in \{1, \ldots, n\} \text{ and every permutation } \underset{\sim}{\pi}. \tag{2.6}$$

Rules satisfying (2.5) have been called "natural" in some of the selection literature (for example, Eaton 1967). Gupta and Miescke (1982) have shown that for problems involving exponential families, selection rules satisfying (2.5) form an essentially complete class among all permutation invariant rules for many loss functions. Both Eaton (1967) and Gupta and Miescke (1982) allow randomization to break ties. The permutation invariance assumption (2.6) is standard and reasonable in light of the permutation invariance of the density g and measure σ.

We have restricted attention to nonrandomized selection rules. Typically, in the fixed subset size selection rules, ties are broken at random. If $F(\underset{\sim}{x}; \underset{\sim}{\lambda})$ is such that ties among the X_i's occur with probability zero, then our results may apply to these fixed subset size rules. But if ties occur with positive probability (for example, $F(\underset{\sim}{x}; \underset{\sim}{\lambda})$ is a multinomial distribution), then our results are not directly applicable to these randomized, fixed subset size rules. But our results are applicable to the random subset size rules for the multinomial such as those considered by Gupta and Nagel (1967) and Gupta and Huang (1975), since the multinomial density is DT.

The following lemma will be used to prove the monotonicity results in the next section.

LEMMA 2.1. If $A = \{a_1, \ldots, a_k\} \subset \{1, \ldots, n\}$ is better than $B = \{b_1, \ldots, b_k\} \subset \{1, \ldots, n\}$, then there exist vectors $\underset{\sim}{\lambda}'$ and $\underset{\sim}{\lambda}''$ such that a) $\{\lambda_{a_1}, \ldots, \lambda_{a_k}\} = \{\lambda'_{n-k+1}, \ldots, \lambda'_n\}$, b) $\{\lambda_{b_1}, \ldots, \lambda_{b_k}\} = \{\lambda''_{n-k+1}, \ldots, \lambda''_n\}$, and c) λ'' is a transposition of $\underset{\sim}{\lambda}'$.

PROOF. We will define $\underset{\sim}{\lambda}'$ and $\underset{\sim}{\lambda}''$ and then show they have the required characteristics. Let $r = |A^c \cap B|$. Note that $r = |A \cap B^c|$, $k - r = |A \cap B|$ and $n-k-r = |A^c \cap B^c|$. Let $\underset{\sim}{\lambda}^1 = (\lambda'_1, \ldots, \lambda'_{n-k-r})$ be the elements of $\{\lambda_i\colon\ i \in A^c \cap B^c\}$ in an arbitrary but fixed order and let $(\lambda''_1, \ldots, \lambda''_{n-k-r}) = \underset{\sim}{\lambda}^1$. Let $\underset{\sim}{\lambda}^2 = (\lambda'_{n-k-r+1}, \ldots, \lambda'_{n-k})$ be the elements of $\{i\colon\ i \in A^c \cap B\}$ arranged so that $\lambda'_{n-k-r+1} \leq \ldots \leq \lambda'_{n-k}$ and let $(\lambda''_{n-k+1}, \ldots, \lambda''_{n-k+r}) = \underset{\sim}{\lambda}^2$. Let $\underset{\sim}{\lambda}^3 = (\lambda'_{n-k+1}, \ldots, \lambda'_{n-k+r})$ be the elements of $\{\lambda_i\colon\ i \in A \cap B^c\}$ arranged so that $\lambda'_{n-k+1} \leq \ldots \leq \lambda'_{n-k+r}$ and let $(\lambda''_{n-k-r+1}, \ldots, \lambda''_{n-k}) = \underset{\sim}{\lambda}^3$. Let $\underset{\sim}{\lambda}^4 = (\lambda'_{n-k+r+1}, \ldots, \lambda'_n)$ be the elements of $\{\lambda_i\colon\ i \in A \cap B\}$ arranged in an arbitrary but fixed order and let $(\lambda''_{n-k+r+1}, \ldots, \lambda''_n) = \underset{\sim}{\lambda}^4$. The two vectors are $\underset{\sim}{\lambda}' = (\underset{\sim}{\lambda}^1, \underset{\sim}{\lambda}^2, \underset{\sim}{\lambda}^3, \underset{\sim}{\lambda}^4)$ and $\underset{\sim}{\lambda}'' = (\underset{\sim}{\lambda}^1, \underset{\sim}{\lambda}^3, \underset{\sim}{\lambda}^2, \underset{\sim}{\lambda}^4)$.

Clearly, a) and b) are true by the definition of $\underset{\sim}{\lambda}'$ and $\underset{\sim}{\lambda}''$. To show that $\underset{\sim}{\lambda}''$ is a transposition of $\underset{\sim}{\lambda}'$, it suffices to show that $\lambda'_{n-k+i} \geq \lambda'_{n-k-r+i}$, $i = 1, \ldots, r$; for, if this is true then $\underset{\sim}{\lambda}''$ can be obtained from $\underset{\sim}{\lambda}'$ by the sequence of r simple transpositions which switch λ'_{n-k+i} and $\lambda'_{n-k-r+i}$, $i = 1, \ldots, r$.

To verify that $\lambda'_{n-k+i} \geq \lambda'_{n-k-r+i}$, $i = 1, \ldots, r$, fix i. Let $s = |A \cap B \cap \{\lambda_j: \lambda_j \geq \lambda'_{n-k-r+i}\}|$. At least $s + r - i + 1$ elements of B are greater than or equal to $\lambda'_{n-k-r+i}$ because the coordinates of $\underset{\sim}{\lambda}^2$ are in nondecreasing order. Since A is better than B, corresponding to each of these there must be an element in A which is greater than or equal to $\lambda'_{n-k-r+i}$. The definition of s implies $|A \cap B^c \cap \{\lambda_j: \lambda_j \geq \lambda'_{n-k-r+i}\}| \geq r - i + 1$. Since the elements of $\underset{\sim}{\lambda}^3$ are in nondecreasing order, $\lambda'_{n-k+j} \geq \lambda'_{n-k-r+i}$, $j = i, \ldots, r$. In particular $\lambda'_{n-k+i} \geq \lambda'_{n-k-r+i}$, as was to be shown.||

3. <u>Monotonicity properties</u>. In this section we prove the monotonicity properties (1.1), (1.2), and (1.3). Lemma 3.1 will be used in these proofs.

<u>LEMMA 3.1</u>. Let $\underset{\sim}{x} \in R^n$, $\underset{\sim}{I} \in \{0, 1\}^n$, and $m \in R$. Suppose $\psi_1, \ldots, \psi_n$, the individual selection probabilities of a selection rule S, satisfy (2.5) and (2.6). Define

$$h_1(\underset{\sim}{I}, \underset{\sim}{x}) = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} I_i \, \psi_i(\underset{\sim}{x}) \geq m \\ 0 & \text{if } \sum_{i=1}^{n} I_i \, \psi_i(\underset{\sim}{x}) < m \end{cases} \tag{3.1}$$

and

$$h_2(\underset{\sim}{I}, \underset{\sim}{x}) = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} (1 - I_i)\psi_i(x) \leq m \\ 0 & \text{if } \sum_{i=1}^{n} (1 - I_i)\psi_i(x) > m . \end{cases} \tag{3.2}$$

Then h_1 and h_2 are DT.

PROOF. We will prove the result for h_1. The proof for h_2 is similar. By 5., page 724 of HPS(1977), it suffices to show that $h(\underset{\sim}{I}, \underset{\sim}{x}) = \sum_{i=1}^{n} I_i \psi_i(\underset{\sim}{x})$ is DT.

The function h is permutation invariant since by (2.6), $h(\underset{\sim}{I} \circ \underset{\sim}{\pi}; \underset{\sim}{x} \circ \underset{\sim}{\pi}) = \sum_{i=1}^{n} I_{\pi_i} \psi_i(\underset{\sim}{x} \circ \underset{\sim}{\pi}) = \sum_{i=1}^{n} I_{\pi_i} \psi_{\pi_i}(\underset{\sim}{x}) = \sum_{i=1}^{n} I_i \; \psi_i(\underset{\sim}{x}) = h(\underset{\sim}{I}, \underset{\sim}{x})$ for every permutation $\underset{\sim}{\pi}$. To verify (2.3), suppose $I_1 = \ldots = I_{r-1} = 0$, $I_r = \ldots = I_n = 1$, and $x_1 \le \ldots \le x_n$. Then $h(\underset{\sim}{I}, \underset{\sim}{x}) = \sum_{s=r}^{n} \psi_s(\underset{\sim}{x})$. For $\underset{\sim}{\pi} >^t \underset{\sim}{\pi}'$ defined by (2.1), (2.5) and (2.6) implies

$$\psi_i(\underset{\sim}{x} \circ \underset{\sim}{\pi}') = \psi_j(\underset{\sim}{x} \circ \underset{\sim}{\pi}) = \psi_{\pi_j}(\underset{\sim}{x}) \ge \psi_{\pi_i}(\underset{\sim}{x}) = \psi_i(\underset{\sim}{x} \circ \underset{\sim}{\pi}) = \psi_j(\underset{\sim}{x} \circ \underset{\sim}{\pi}')$$

and

$$\psi_s(\underset{\sim}{x} \circ \underset{\sim}{\pi}) = \psi_s(\underset{\sim}{x} \circ \underset{\sim}{\pi}') \quad \text{for } s = 1, \ldots, n,\ s \ne i, j.$$

Thus, for $i < j \le r - 1$ or $r \le i < j$, $h(\underset{\sim}{I}, \underset{\sim}{x} \circ \underset{\sim}{\pi}) = h(\underset{\sim}{I}, \underset{\sim}{x} \circ \underset{\sim}{\pi}')$, and for $i \le r - 1 < r \le j$, $h(\underset{\sim}{I}, \underset{\sim}{x} \circ \underset{\sim}{\pi}) \ge h(\underset{\sim}{I}, \underset{\sim}{x} \circ \underset{\sim}{\pi}')$. Therefore, h is DT. ||

THEOREM 3.1. Suppose the density $g(\underset{\sim}{x}; \underset{\sim}{\lambda})$ is DT. Suppose the individual selection probabilities of the selection rule S satisfy (2.5) and (2.6). Let $A \subset \{1, \ldots, n\}$ and $B \subset \{1, \ldots, n\}$. If A is better than B, then

$$P_{\underset{\sim}{\lambda}}(|A \cap S(\underset{\sim}{X})| \ge m) \ge P_{\underset{\sim}{\lambda}}(|B \cap S(\underset{\sim}{X})| \ge m) \quad \text{for every } m \in R. \tag{3.3}$$

PROOF. Let h_1 be defined by (3.1). Then by the Composition

Theorem 3.3 of HPS(1977), $H_1(\underset{\sim}{I}, \underset{\sim}{\lambda}) = \int h_1(\underset{\sim}{I}, \underset{\sim}{x}) g(\underset{\sim}{x}; \underset{\sim}{\lambda}) d\sigma(\underset{\sim}{x})$ is a DT function. For any $C \subset \{1, \ldots, n\}$, define $\underset{\sim}{I}^C = (I_1^C, \ldots, I_n^C)$ by $I_i^C = 1$ if $i \in C$ and $I_i^C = 0$ if $i \notin C$. Then it follows that

$$h_1(\underset{\sim}{I}^C, \underset{\sim}{x}) = \begin{cases} 1 & \text{if } |C \cap S(\underset{\sim}{x})| \geq m \\ 0 & \text{if } |C \cap S(\underset{\sim}{x})| < m \end{cases} .$$

Thus $P_{\underset{\sim}{\lambda}}(|C \cap S(\underset{\sim}{X})| \geq m) = H_1(\underset{\sim}{I}^C, \underset{\sim}{\lambda})$.

Let $\underset{\sim}{\pi}'(\underset{\sim}{\pi}'')$ be the permutation such that $\underset{\sim}{\lambda} \circ \underset{\sim}{\pi}'(\underset{\sim}{\lambda} \circ \underset{\sim}{\pi}'') = \underset{\sim}{\lambda}'(\lambda'')$ where $\underset{\sim}{\lambda}'(\underset{\sim}{\lambda}'')$ is defined by Lemma 2.1. Then $\underset{\sim}{I}^A \circ \underset{\sim}{\pi}' = I^B \circ \underset{\sim}{\pi}''$ $= (0, \ldots, 0, 1, \ldots, 1)$, a vector of $n - k$ zeros followed by k ones. Since $\underset{\sim}{\lambda}''$ is a transposition of $\underset{\sim}{\lambda}'$ and H_1 is DT, we obtain

$$\begin{aligned} P_{\underset{\sim}{\lambda}}(|A \cap S(\underset{\sim}{X})| \geq m) &= H_1(\underset{\sim}{I}^A, \underset{\sim}{\lambda}) \\ &= H_1(\underset{\sim}{I}^A \circ \underset{\sim}{\pi}', \underset{\sim}{\lambda} \circ \underset{\sim}{\pi}') \\ &= H_1(\underset{\sim}{I}^B \circ \underset{\sim}{\pi}'', \underset{\sim}{\lambda} \circ \underset{\sim}{\pi}') \\ &\geq H_1(\underset{\sim}{I}^B \circ \underset{\sim}{\pi}'', \underset{\sim}{\lambda} \circ \underset{\sim}{\pi}'') \\ &= H_1(\underset{\sim}{I}^B, \underset{\sim}{\lambda}) \\ &= P_{\underset{\sim}{\lambda}}(|B \cap S(\underset{\sim}{X})| \geq m). \quad \| \end{aligned} \tag{3.4}$$

The conclusion of Theorem 3.1 can be restated as, "$|A \cap S(\underset{\sim}{X})|$ is stochastically larger than $|B \cap S(\underset{\sim}{X})|$." This implies other inequalities in addition to (3.3) such as $E_{\underset{\sim}{\lambda}}|A \cap S(\underset{\sim}{X})| \geq E_{\underset{\sim}{\lambda}}|B \cap S(\underset{\sim}{X})|$.

<u>THEOREM 3.2</u>. Under the assumptions of Theorem 3.1,

$$P_{\underset{\sim}{\lambda}}(|A^c \cap S(\underset{\sim}{X})| \leq m) \geq P_{\underset{\sim}{\lambda}}(|B^c \cap S(\underset{\sim}{X})| \leq m) \text{ for every } m \in R. \tag{3.5}$$

PROOF. Let h_2 be defined by (3.2). By the Composition Theorem, $H_2(\underset{\sim}{I}, \underset{\sim}{\lambda}) = \int h_2(\underset{\sim}{I}, \underset{\sim}{x})g(\underset{\sim}{x}; \underset{\sim}{\lambda})d\sigma(\underset{\sim}{x})$ is DT. But for I^C defined as in Theorem 3.1, $H_2(\underset{\sim}{I}^C, \underset{\sim}{\lambda}) = P_{\underset{\sim}{\lambda}}(|C^c \cap S(\underset{\sim}{X})| \le m)$. Thus, arguing as in (3.4), we obtain (3.5). ||

THEOREM 3.3. Under the assumptions of Theorem 3.1,

$$P_{\underset{\sim}{\lambda}}(A = S(\underset{\sim}{X})) \ge P_{\underset{\sim}{\lambda}}(B = S(\underset{\sim}{X})). \tag{3.6}$$

PROOF. Let $k = |A| = |B|$. Let $h(\underset{\sim}{I}, \underset{\sim}{x}) = h_1(\underset{\sim}{I}, \underset{\sim}{x})h_2(\underset{\sim}{I}, \underset{\sim}{x})$, where h_1 is defined in (3.1) with $m = k$ and h_2 is defined in (3.2) with $m = 0$. By Theorem 3.6 of HPS(1977), h is a DT function. Thus $H(\underset{\sim}{I}, \underset{\sim}{\lambda}) = \int h(\underset{\sim}{I}, \underset{\sim}{x})g(\underset{\sim}{x}; \underset{\sim}{\lambda})d\sigma(\underset{\sim}{x})$ is DT by the Composition Theorem. Furthermore, for $\underset{\sim}{I}^C$ defined as in Theorem 3.1, $H(\underset{\sim}{I}^C, \underset{\sim}{\lambda}) = P_{\underset{\sim}{\lambda}}(C = S(\underset{\sim}{X}))$. Thus, arguing as in (3.4), we obtain (3.6). ||

4. Additional parameters and statistics.

In many problems, the distribution of the observations depends not only on $\underset{\sim}{\lambda}$, the parameter of interest, but also on another parameter $\underset{\sim}{\nu}$. In many problems the selection rule depends not only on $\underset{\sim}{X}$ but also on another statistic $\underset{\sim}{Y}$. In this section we present conditions in this more general framework under which the monotonicity properties (1.1), (1.2), and (1.3) hold.

As an example of this type of problem, consider the selection of the largest normal mean. Suppose $X_1, \ldots, X_n$ are independent normal means from normal populations. The mean of X_i is λ_i and all the X_i have the same variance ν. Suppose that Y is an estimate of ν which is independent of $\underset{\sim}{X}$ and such that $r\frac{Y}{\nu}$ has a χ^2 distribution with r degrees of freedom. Then Gupta (1956) proposed the following selection rule to select a random size subset including the largest λ_i:

Include i in the selected subset

$S(\underset{\sim}{X}, \underset{\sim}{Y})$ if and only if

$$X_i \geq \max_{1 \leq j \leq n} X_j - d\sqrt{Y},$$

where d is a constant chosen by the experimenter.

Our assumptions will imply that this selection rule satisfies the monotonicity properties (1.1), (1.2), and (1.3).

We assume that the observation $(\underset{\sim}{X}, \underset{\sim}{Y})$ has a density $g(\underset{\sim}{x},\underset{\sim}{y};\underset{\sim}{\lambda},\underset{\sim}{\nu})$ with respect to a measure $\sigma(\underset{\sim}{x}) \times \mu(\underset{\sim}{y})$ where σ satisfies $\int_A d\sigma(\underset{\sim}{x}) = \int_A d\ (\underset{\sim}{x} \cdot \underset{\sim}{\pi})$ for each permutation $\underset{\sim}{\pi}$ and Borel set $A \subset R^n$. We assume that for each fixed $\underset{\sim}{y} \in \mathcal{Y}$, the sample space of $\underset{\sim}{Y}$, and each fixed $\underset{\sim}{\nu} \in N$, the set of possible values of $\underset{\sim}{\nu}$, $g(\underset{\sim}{x}, \underset{\sim}{y}; \underset{\sim}{\lambda}, \underset{\sim}{\nu})$ is a DT function of $\underset{\sim}{x}$ and $\underset{\sim}{\lambda}$.

Let $\psi_1(\underset{\sim}{x}, \underset{\sim}{y}), \ldots, \psi_n(\underset{\sim}{x}, \underset{\sim}{y})$ denote the individual selection probabilities of a nonrandomized selection rule $S(\underset{\sim}{X}, \underset{\sim}{Y})$. We assume that for every $\underset{\sim}{y} \in \mathcal{Y}$, if $\psi_i(\underset{\sim}{x}, \underset{\sim}{y}) = 1$ and $x_j \geq x_i$ then $\psi_j(\underset{\sim}{x}, \underset{\sim}{y}) = 1$. We also assume that $\underset{\sim}{y} \in \mathcal{Y}$, $\underset{\sim}{x} \in R^n$, $i \in \{1, \ldots, n\}$, and $\underset{\sim}{\pi}$ a permutation imply $\psi_{\pi_i}(\underset{\sim}{x}, \underset{\sim}{y}) = \psi_i(\underset{\sim}{x} \cdot \underset{\sim}{\pi}, \underset{\sim}{y})$.

<u>THEOREM 4.1</u>. Suppose the above assumptions concerning the distribution of $(\underset{\sim}{X}, \underset{\sim}{Y})$ and the form of $S(\underset{\sim}{X}, \underset{\sim}{Y})$ hold. Let $A \subset \{1, \ldots, n\}$ and $B \subset \{1, \ldots, n\}$. If A is better than B, then

$$P_{\underset{\sim}{\lambda},\underset{\sim}{\nu}}(|A \cap S(\underset{\sim}{X},\underset{\sim}{Y})| \geq m) \geq P_{\underset{\sim}{\lambda},\underset{\sim}{\nu}}(|B \cap S(\underset{\sim}{X},\underset{\sim}{Y})| \geq m) \text{ for every } m \in R, \quad (4.1)$$

$$P_{\underset{\sim}{\lambda},\underset{\sim}{\nu}}(|A^c \cap S(\underset{\sim}{X},\underset{\sim}{Y})| \leq m) \geq P_{\underset{\sim}{\lambda},\underset{\sim}{\nu}}(|B^c \cap S(\underset{\sim}{X},\underset{\sim}{Y})| \leq m) \text{ for every } m \in R, \quad (4.2)$$

and

$$P_{\underset{\sim}{\lambda},\underset{\sim}{\nu}}(A = S(\underset{\sim}{X}, \underset{\sim}{Y})) \geq P_{\underset{\sim}{\lambda},\underset{\sim}{\nu}}(B = S(\underset{\sim}{X}, \underset{\sim}{Y})). \quad (4.3)$$

PROOF. We will outline the proof of (4.1). The proofs of (4.2) and (4.3) are similar.

For every $\underset{\sim}{I} \in \{0, 1\}^n$, $\underset{\sim}{x} \in R^n$, and $\underset{\sim}{y} \in \mathcal{Y}$, define

$$h(\underset{\sim}{I}, \underset{\sim}{x}, \underset{\sim}{y}) = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} I_i \psi_i(\underset{\sim}{x}, \underset{\sim}{y}) \geq m \\ 0 & \text{if } \sum_{i=1}^{n} I_i \psi_i(\underset{\sim}{x}, \underset{\sim}{y}) < m . \end{cases}$$

Arguing as in Lemma 3.1, we can verify that for each fixed $\underset{\sim}{y} \in \mathcal{Y}$, h is a DT function of $\underset{\sim}{I}$ and $\underset{\sim}{x}$. Thus by the Composition Theorem, $H_{\underset{\sim}{y},\underset{\sim}{\nu}}(\underset{\sim}{I}, \underset{\sim}{\lambda}) = \int h(\underset{\sim}{I}, \underset{\sim}{x}, \underset{\sim}{y}) g(\underset{\sim}{x}, \underset{\sim}{y}; \underset{\sim}{\lambda}, \underset{\sim}{\nu}) d\sigma(\underset{\sim}{x})$ is a DT function of $\underset{\sim}{I}$ and $\underset{\sim}{\lambda}$ for each fixed $\underset{\sim}{y} \in \mathcal{Y}$ and $\underset{\sim}{\nu} \in N$. Consequently, by Theorem 3.2 of HPS(1977), $H_{\underset{\sim}{\nu}}(\underset{\sim}{I}, \underset{\sim}{\lambda}) = \int H_{\underset{\sim}{y},\underset{\sim}{\nu}}(\underset{\sim}{I}, \underset{\sim}{\lambda}) d\mu(\underset{\sim}{y})$ is a DT function of $\underset{\sim}{I}$ and $\underset{\sim}{\lambda}$ for each fixed $\underset{\sim}{\nu} \in N$. But for $\underset{\sim}{I}^C$ defined as in Theorem 3.1, $H_{\underset{\sim}{\nu}}(\underset{\sim}{I}^C, \underset{\sim}{\lambda}) = P_{\underset{\sim}{\lambda},\underset{\sim}{\nu}}(|C \cap S(\underset{\sim}{X}, \underset{\sim}{Y})| \geq m)$. Thus, arguing as in (3.4), we obtain (4.1). ||

As another example of the generality of this result, consider the comparison with a standard problem presented by Gupta and Sobel (1958). Let $X_0, X_1, \ldots, X_n$ denote independent sample means from normal populations. The mean of X_i is λ_i, $i = 0, 1, \ldots, n$. The variance of X_0 is σ^2/m and the variance of X_i is σ^2/r, $i = 1, \ldots, n$. The parameters $\lambda_0, \lambda_1, \ldots, \lambda_n$ and σ^2 are unknown but m and r are known (sample sizes). Let S^2 be independent of $(X_0, X_1, \ldots, X_n)$ and such that vS^2/σ^2 has a χ^2 distribution with v degrees of freedom. The goal is to select a subset of $\{1, \ldots, n\}$ which contains $\{i: \lambda_i > \lambda_0\}$. Gupta and Sobel (1958) proposed the following selection rule:

> Include i in the selected subset $S(\underset{\sim}{X}, \underset{\sim}{Y})$ if and only if
>
> $$X_i \geq X_0 - d\sqrt{S^2/r}$$
>
> where d is a constant chosen by the experimenter.

With $\underset{\sim}{X} = (X_1, \ldots, X_n)$, $\underset{\sim}{\lambda} = (\lambda_1, \ldots, \lambda_n)$, $\underset{\sim}{Y} = (X_0, S^2)$, and $\underset{\sim}{\nu} = (\lambda_0, \sigma^2)$, the assumptions of Theorem 4.1 are readily verified for Gupta and Sobel's selection rule. These authors considered the more general problem in which the r's (sample sizes) associated with the various X_i's are different. The X_i's must all have the same r for the assumptions of Theorem 4.1 to be satisfied. To our knowledge, monotonicity properties such as these have not previously been considered in the comparison with a standard problem. But they are as desirable in this framework as in other formulations (Bechhofer, 1954, Gupta, 1965) of the selection problem.

REFERENCES

Bechhofer, R.E. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances. Ann. Math. Statist. 25 16-39.

Eaton, M.L. (1967). Some optimum properties of ranking procedures. Ann. Math. Statist. 38 124-137.

Gupta, S.S. (1956). On a decision rule for a problem in ranking means. Ph.D. Thesis (Mimeo. Ser. No. 150). Inst. of Statist., Univ. of North Carolina, Chapel Hill.

Gupta, S.S. (1965). On some multiple decision (selection and ranking) rules. Technometrics 7 225-245.

Gupta, S.S. and Huang, D.Y. (1975). On subset selection procedures for Poisson populations and some applications to the multinomial selection problems. Applied Statistics (Ed. R.P. Gupta), North Holland, Amsterdam, pp. 97-109.

Gupta, S.S. and Miescke, K.J. (1982). Sequential selection procedures — a decision theoretic approach. Tech. Report 82-6. Dept. of Statist., Purdue Univ., West Lafayette, Indiana.

Gupta, S.S. and Nagel, K. (1967). On selection and ranking procedures and order statistics from the multinomial distribution. Sankhyā Ser. B 29 1-34.

Gupta, S.S. and Sobel, M. (1958). On selecting a subset which contains all populations better than a standard. Ann. Math. Statist. 29 235-244.

Hollander, M., Proschan, F., and Sethuraman, J. (1977). Functions decreasing in transposition and their applications in ranking problems. Ann. Statist. 5 722-733.

Hsu, J.C. (1977). On some decision-theoretic contributions to the problem of subset selection. Ph.D. Thesis (Mimeo Ser. No. 491). Dept. of Statist., Purdue Univ., West Lafayette, Indiana.

Lehmann, E.L. (1961). Some Model I problems of selection. Ann. Math. Statist. 32 990-1012.

Marshall, A.W. and Olkin, I. (1979). Inequalities: Theory of Majorization and Its Applications. Academic Press, New York.

Santner, T.J. (1975). A restricted subset selection approach to ranking and selection problems. Ann. Statist. 3 334-349.

Tong, Y.L. (1969). On partitioning a set of normal populations by their locations with respect to a control. Ann. Math. Statist. 40 1300-1324.

CURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1. REPORT NUMBER FSU M647 USARO D-57 AFOSR 82-154	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and subtitle) Monotonicity in Selection Problems: A Unified Approach		5. TYPE OF REPORT & PERIOD COVERED Technical 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Roger L. Berger and Frank Proschan		8. CONTRACT OR GRANT NUMBER(s) USARO DAAG 29-82-K-0168 AFOSR 82-K-0007
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Florida State University Tallahassee, FL 32306		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Air Force Air Force Office of Scientific Research Bolling Air Force Base, D.C. 20332 U.S. Army Research Office - Durham P.O. Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE September, 1982 13. NUMBER OF PAGES 15 15. SECURITY CLASS. (of this report) 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
14. MONITORING OFFICE NAME AND ADDRESS (if different from controlling office)		

16. DISTRIBUTION STATEMENT (of this report)

Distribution unlimited

s prime

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from report)

Lambda (approximation)

18. SUPPLEMENTARY NOTES

THE VIEW, OPINIONS, AND/OR FINDINGS CONTAINED IN THIS REPORT ARE THOSE OF THE AUTHOR(S) AND SHOULD NOT BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, POLICY, OR DECISION, UNLESS SO DESIGNATED BY OTHER DOCUMENTATION.

Lambda (approximation)

lambda

19. KEY WORDS

Ranking and selection, decreasing in transposition, monotonicity, arrangement increasing, comparison with a standard

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Let $\underset{\sim}{X} = (X_1^{\Pi}, \ldots, X_n^{\Pi})$ have a density $g(\underset{\sim}{x}, \underset{\sim}{\lambda})$ which is decreasing in transposition, where $\underset{\sim}{\lambda} = (\lambda_1^{\Pi}, \ldots, \lambda_n^{\Pi})$. Suppose one wishes to select a subset of $\{1, \ldots, n\}$ containing the subscripts associated with the largest values of the λ_i^{Π}'s. Let $S(\underset{\sim}{x})$ be a permutation invariant selection rule which always selects a subset associated with the largest values of the X_i's. Let $A = \{i(1),\ldots,i(r)\} \subset \{1,\ldots,n\}$ and $B = \{j(1),\ldots,j(r)\} \subset \{1,\ldots,n\}$ be such that $\lambda_{i(s)} \geq \lambda_{j(s)}$, $s = 1,\ldots,r$. Then the following three inequalities are proved. ($|C|$ denotes the number of elements in C. C^c denotes the complement of C.) (i) $P_{\underset{\sim}{\lambda}}(|A \cap S(\underset{\sim}{X})| \geq m) \geq P_{\underset{\sim}{\lambda}}(|B \cap S(\underset{\sim}{X})| \geq m)$ for every $m \in R$, (ii) $P_{\underset{\sim}{\lambda}}(A = S(\underset{\sim}{X})) \geq P_{\underset{\sim}{\lambda}}(B = S(\underset{\sim}{X}))$, and (iii) $P_{\underset{\sim}{\lambda}}(|A^c \cap S(\underset{\sim}{X})| \leq m) \geq P_{\underset{\sim}{\lambda}}(|B^c \cap S(\underset{\sim}{X})| \leq m)$ for every $m \in R$. These generalized monotonicity properties are derived using a unified approach. The results apply to selection rules proposed under several formulations of the selection problem.

LME

-8